

## CPSARST – An efficient circular permutation search tool applied to the detection of novel protein structural relationships

Wei-Cheng Lo and Ping-Chiang Lyu

### Additional data file 5 – Statistics of protein structural database searches with broad criteria

Description: The criteria of these database searches are  $\text{RMSD} \leq 6 \text{ \AA}$ ,  $\text{CP score} \leq 0.1$ , and permutation size  $\geq 10\%$ , which are used for a broader estimation of the natural prevalence of circular permutation.

Database		nrPDB-90	nrSCOP-90	
No. of proteins		14422	11688	
No. of candidate pairs	1. Detected by amino acid sequence	8482	2917	
	2. Detected only by Ramachandran string	313383	249153	
	3. Confirmed after the refinement stage	Total	14067	21809
		Symmetric CP	8656	13896
		Symmetric CP with sequence identity > 30	129	102